AC

AMS 7 10/20/09

PROBABILITY MODELS FOR SUMS & MEANS

(R-63) Your net gain after 1000 \$1 bets on a single number is like:

the sum of 1000 IID draws from this population

long run SD = standard error of S = SE = $SE_{IID}(S)$

IMAGINARY DATA SET
(POSSIBLE S of 1000, 1\$ plays)

$$\begin{bmatrix} -64 \\ -28 \\ \vdots \end{bmatrix}$$

long run mean EV of S: -53\$

Possible variables for SE:		
N	x	
μ	x	
σ	↑ SE ↑	✓
n	↑ SE(S) ↑	✓

MATH FACT

$$SE_{IID}(S) = \frac{\sigma\sqrt{n}}{1} = \sigma\sqrt{n} \Rightarrow \left(\begin{smallmatrix}\text{pop}\\ \text{SD}\end{smallmatrix}\right)\sqrt{\#\text{ of draws}}$$

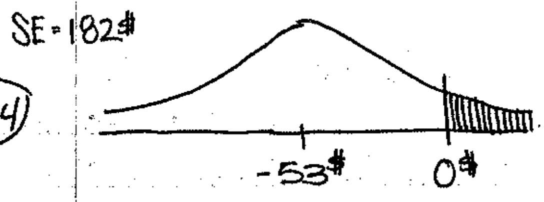
$$SE(S) = \$5.76\sqrt{1000} = 182\$$$

After 1000, 1\$ plays of a single # I expect to be behind by about 53\$ (EV) ± 182\$

- Note: possible range = (-1000\$ → +35,000\$)

long run Histogram of S

(C-124)



← CENTRAL LIMIT THEORUM (CLT) →

① as long as (n) is large, the long run histogram of S (or mean) of (n) IID draws from a pop will look a lot like the normal curve.
Ex. of CLT (R-64)

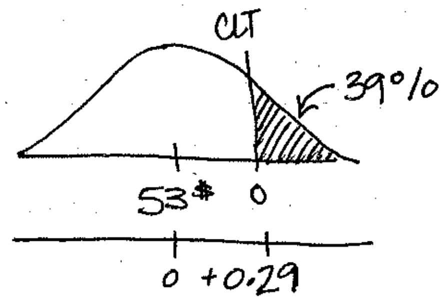
② the more non-normal a pop, the greater (n) needs to be to get a normal curve

③ If the population is normal to begin with, any sample ⓝ will alway yield a normal curve

Would ⓝ = 1000 be large enough?

$$\frac{0\$ - (-53)}{182} = +0.29 = z \text{ value}$$

table → 39%



(σ)

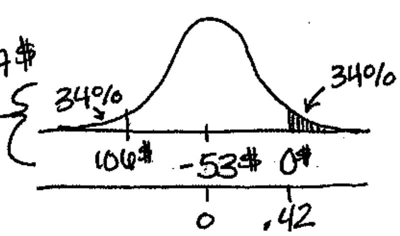
∘ SPLIT: loose 0.05$ ± 4.02$

EV of $ = −53$

long run SD: SE of $ = 127$

Long Run Histogram of $ (split)

SE 127$



→ equal chance of loss/gain

$$\$\frac{0 - (-.53)}{127\$} = 0.42 \rightarrow 34\%$$

Single #: Risk seeking ← Bold play when odds are stacked against you

Split: more risk adverse

CASE STUDY: HYPOKALEMIA (low K^+)

BLOCK OF BUTTER (oz.)

$$\begin{bmatrix} 16 \\ 16 \\ 16 \\ 16 \end{bmatrix} \quad \begin{bmatrix} 16.0 \\ 16.0 \\ 16.0 \\ 16.0 \end{bmatrix} \quad \begin{bmatrix} 15.97 \\ 16.01 \\ \vdots \\ 15.99 \end{bmatrix}$$

↑ ↑ DETERMINISTIC ↑ STOCHASTIC (random)

there is always a level of detail at which a measurement to switch from DETERMINISTIC to STOCHASTIC

AC

○ BASIC MEASUREMENT ERROR MODEL ⇒

$$\text{observation } 1 = (\text{true value}) + (\text{bias}) + \text{“Random ERROR 1”}$$
$$\text{obs } 2 = (\text{true value}) + (\text{bias}) + \text{“Random ERROR 2”}$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$\underbrace{\text{obs } n}_{\text{observable}} = \underbrace{(\text{true value}) + (\text{bias}) + \text{“Random ERROR } n\text{”}}_{\text{Not observable}}$$

IID, mean = 0

good data-gathering activity: unbiased (bias = 0)

Suppose k^+ value truth = 3.8 (not Hypokalemic)

* measuring unbiased

$$3.9 = 3.8 + 0 + (+0.1)$$
$$3.5 = 3.8 + 0 + (-0.3)$$
$$\vdots$$
$$4.0 = 3.8 + 0 + (+0.2)$$

$$\text{mean of } n \text{ obs.} = 3.8 + 0 + \frac{(+0.1) + (-0.3) + \dots + (+0.2)}{n}$$

· Cancellation of ⊕, ⊖ #'s mean of (n) IID Random ERROR (each w/ mean 0) will likely be a lot closer to 0 than any of the errors themselves.

POP	SAMPLE	IMAGINARY DATA
Conceptual: all possible measurements	observed k^+ measurements	
	unbiased, truth = 3.8	

POTTASSIUM

$$N = \infty \quad \begin{bmatrix} \\ \\ \\ \end{bmatrix} \quad \xrightarrow[\text{SRS}]{\text{like}} \quad \text{POTTASSIUM} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} n = 4$$

mean μ = 3.8

SD σ = 0.2

mean $\bar{Y}$

P(misdiagnosis w/ n = 1) =

POP HIST. = hist of 1 reading